**Project Name** - Covoid-19 Monitor

A website to get latest covid-19 related data using the data obtained to get visual analysis tools like graphs and charts.

I will crawl the government websites for getting the datas and updates.This

update will be shown on our website within 24 hrs.This data will be updated

automatically.

**Input:**

Regular text data from Twitter and social media, and checking whether fake

news or not

Ministry of Health and Govt of India and regional Governments data

WHO, CDC

**Output of the system**

1. It will show the total number of affected cases, total deaths cases, total

cured/recovered cases.

2. It will show a graph which will predict the number of affected cases in near future**.**

3. It will show a graph which will predict the number of affected cases in near future

for each and every state in Country individually.

4. It will show the total number of affected cases, total deaths cases, total

cured/recovered cases for each and every state in the country individually.

5. It will show a graph for trend and seasonality analysis.

6. This graph does other statistical calculations like mean, variation.

7. It describes Corona Virus, awareness, Symptoms.

**Intended Audience**

● Data Scientists

● Statistician

● Governments (for prediction based resource distribution)

● Anyone who wants to visualize the situation and future prediction

**Technology Stack**

**Languages**

● HTML

● CSS

● Python

● Java Script

**Libraries**

● Flask

● Matplotlib

● Scikit

● Pandas

**FOR DETAILED CODE AND FORKED DATA VISIT MY GITHUB ACCOUNT:**

**https://github.com/ankit-kumar04**

**Introduction**

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The disease was first identified in 2019 in Wuhan, China, and has since spread globally, resulting in the 2019–20 coronavirus pandemic. Common symptoms include fever, cough and shortness of breath. Muscle pain, sputum production and sore throat are less common. The rate of deaths per number of diagnosed cases is on average 3.4%, ranging from 0.2% in those less than 20 to approximately 15% in those over 80 years old.

**WORKING:**

So Firstly,I will inspect the data, represent it on chart and model future trends with Python, using some open source data science libraries such as Pandas, Matplotlib and Scikit-learn.

**Installing necessary libraries**

pip install pandas matplotlib scikit-learn

**Importing data and extracting fields**

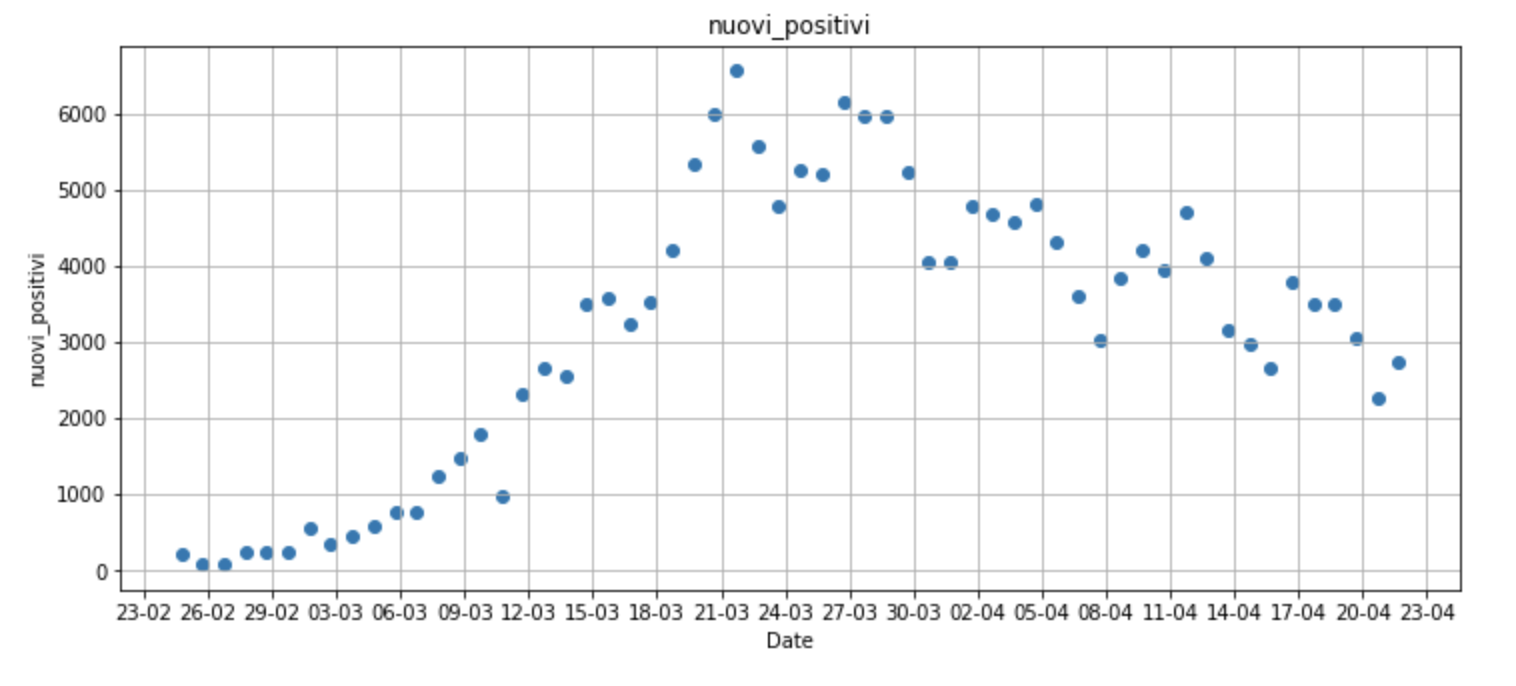
First, I import all the necessary libraries, then I import the COVID dataset GitHub account and we store it in a *Pandas* *Data Frame*. Then, we explore the structure of the table, in order to have a clearer view on the variables.I am taking Italy Data for this project Work.

import pandas as pd  
from datetime import datetime, timedelta  
import matplotlib.pyplot as plt  
from matplotlib.dates import DateFormatter  
import matplotlib.dates as mdatesdata = pd.read\_csv(“<https://raw.githubusercontent.com/pcm-dpc/COVID-19/master/dati-andamento-nazionale/dpc-covid19-ita-andamento-nazionale.csv>")

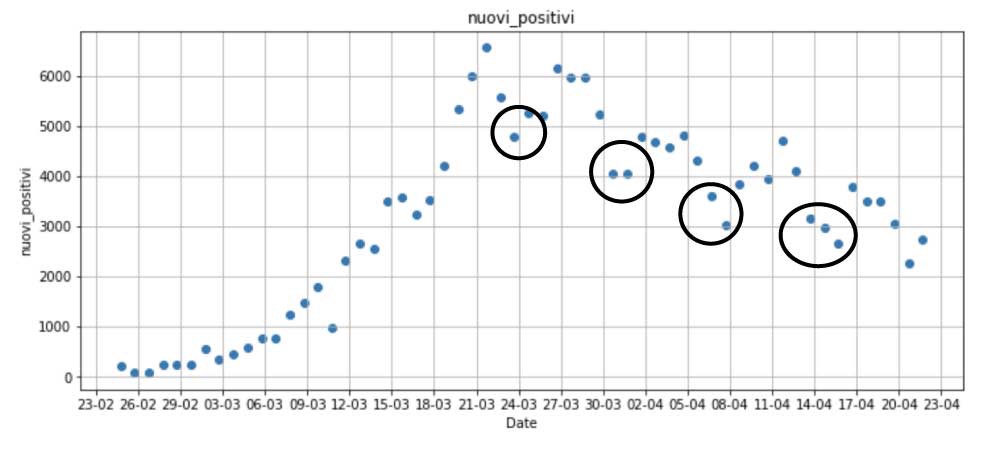
print (data.columns)

In order to calculate the daily increase of those two variables, we run a simple pandas function.

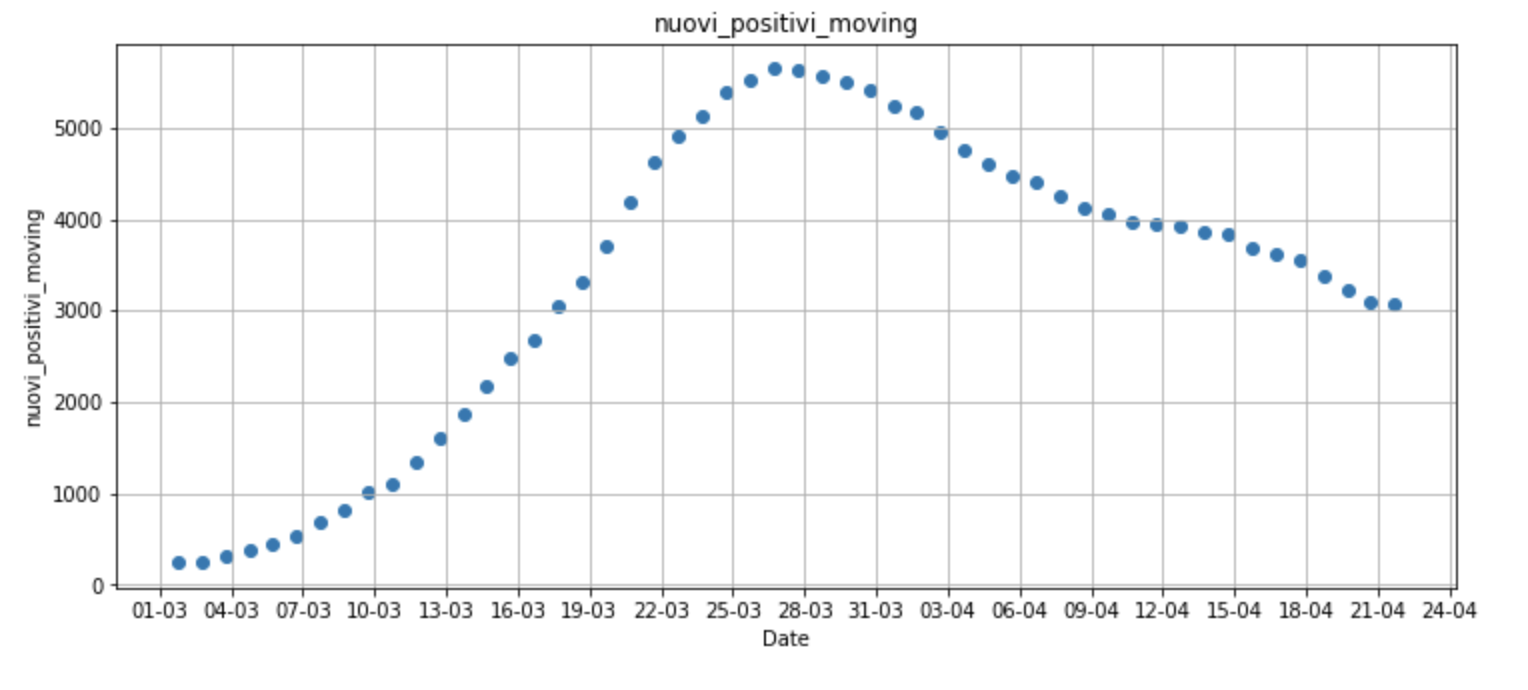
At first glance, seems to be the most important for estimating the epidemic progression: the number of daily new positive cases.



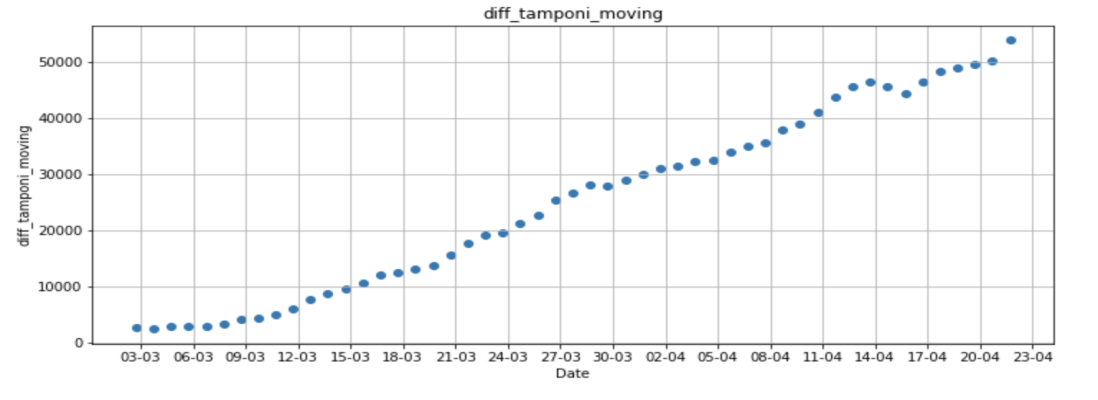
There is a recurrent pattern in the plot above



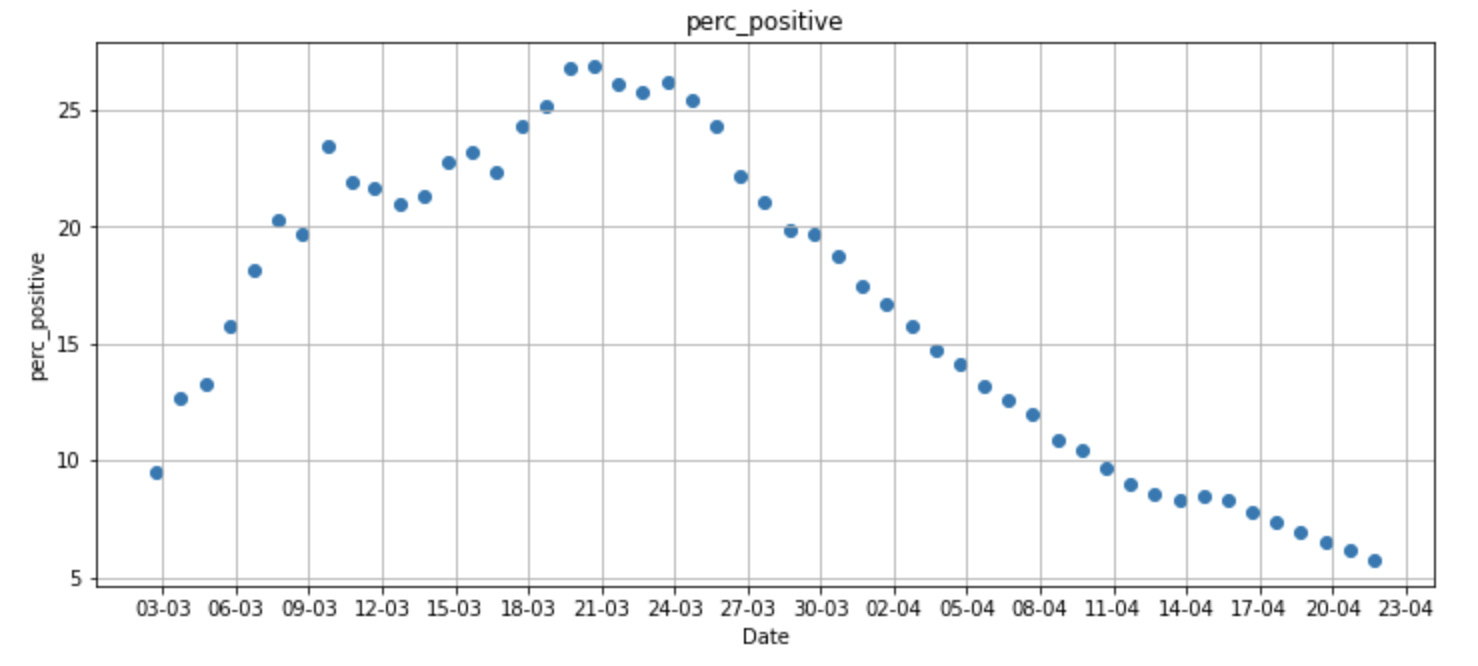
The first moving average is calculated by averaging the first subset of 7 days, and then the subset is changed by moving forward to the next fixed subset, and so on.



The chart above represents the trend of new cases averaged over a 7-days period, masking out weekend anomalies. Nevertheless, the effect of the daily number of swabs tested every day is not fully compensated yet. Clearly, the number of positive cases is strictly correlated with the quantity of tests performed. Let’s take a look to the trend of daily tests performed in Italy …

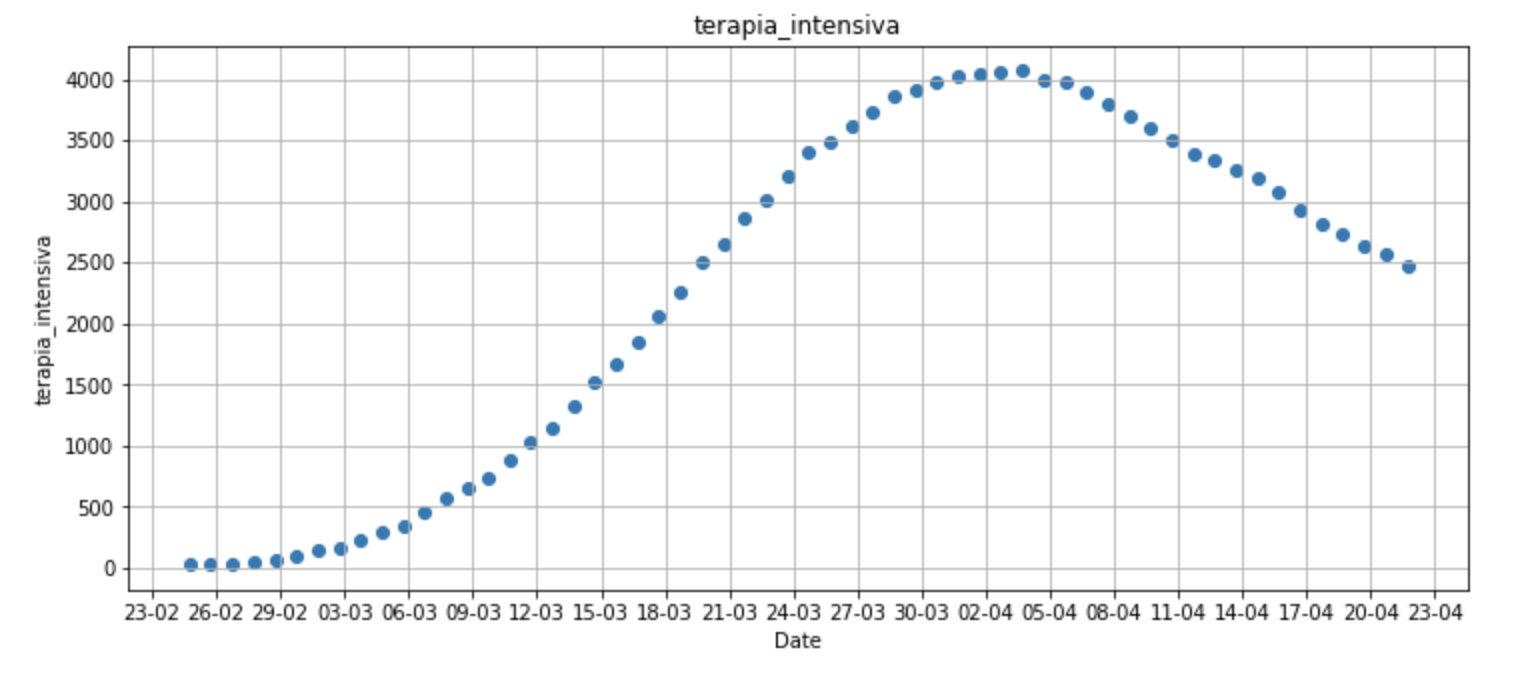


The number of daily swabs is about 20 times higher than the beginning, therefore the variable new pos is suffering from this important bias. In order to find a more representing trend, we now calculate the percentage of new positive over the total daily tests, and we inspect the variation over time.

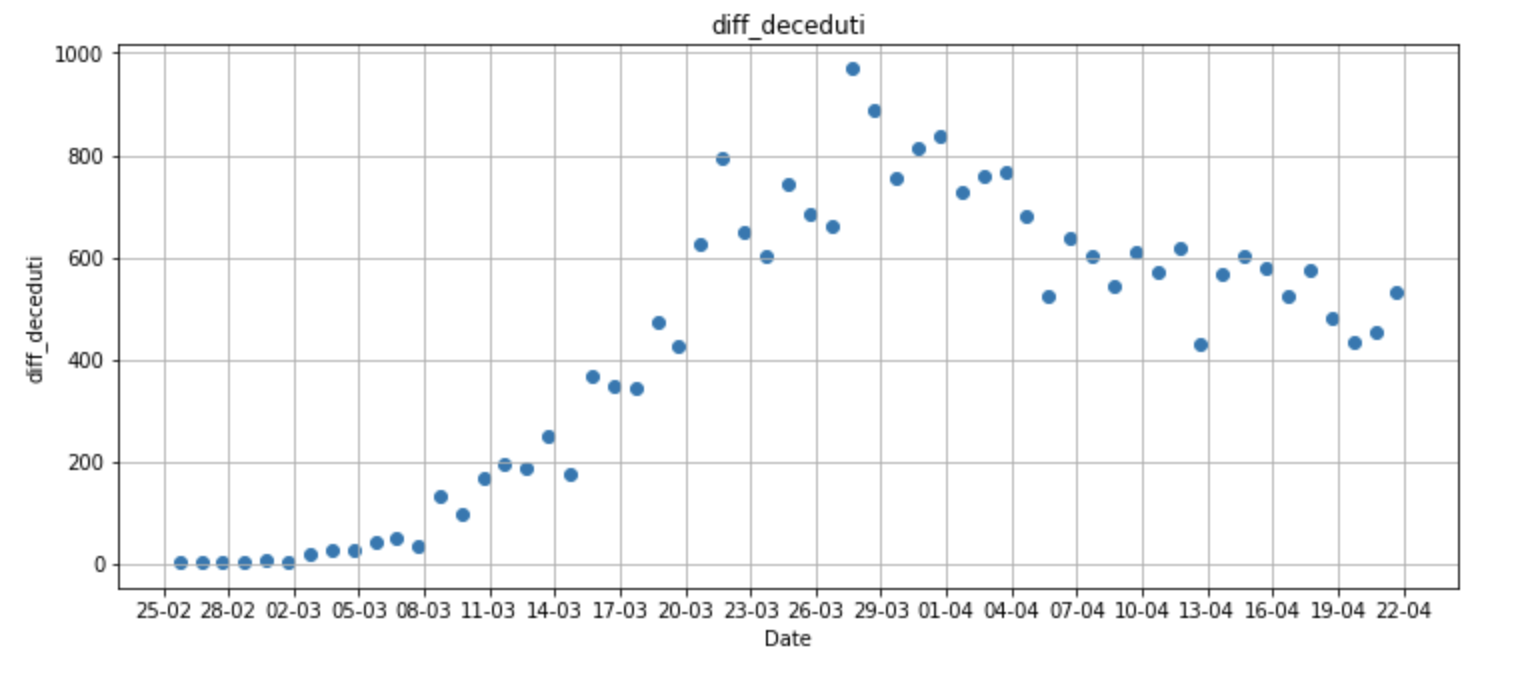


The derived variable per positive provides a more reliable representation of the epidemic progression,

There are other important fields that we should further inspect intensive care and daily deaths, respectively.



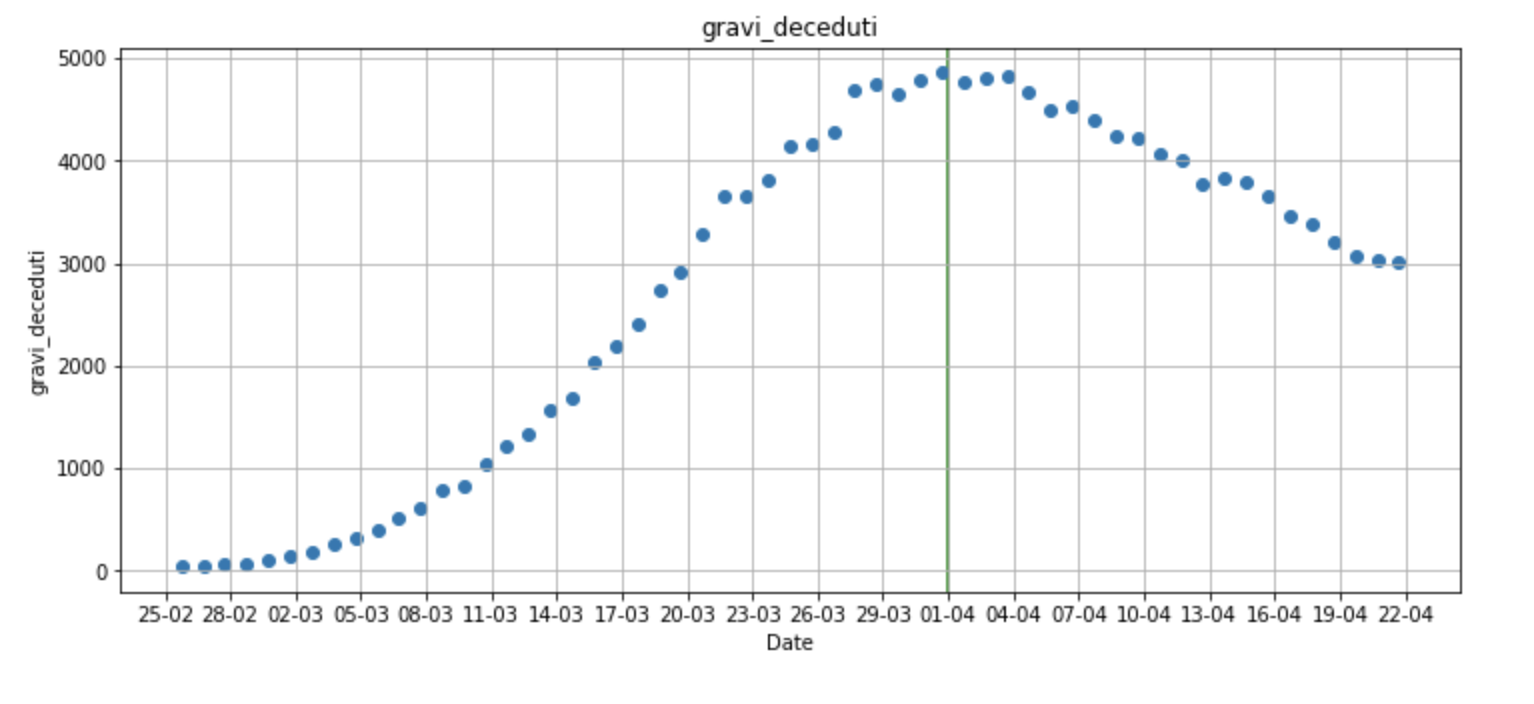
As shown in the chart above, the number of patients currently in intensive care seems to follow a more regular trend.



The above charts shows that daily deaths have been increasing until march 28th, and then it started decreasing at a slower pace.

**Creating derived variable for the project**

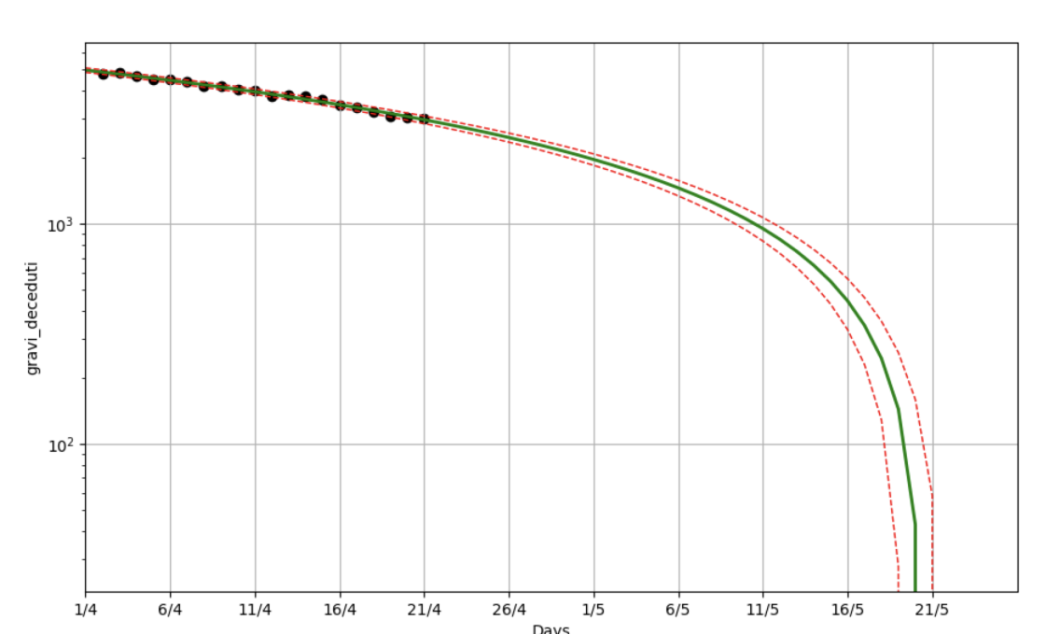
The main assumption is that the combined value of intensive care and daily deaths can be a reliable variable for estimating the current epidemic progression, and for modeling future trends. calculate the sum of patients in severe distress with daily deaths, and plot the resulting values.



**Modeling the epidemic trend**

A Linear Regression model and train it with the data starting from that date, April 1st.Linear Regression is one of the most popular classical machine learning algorithms for supervised learning.

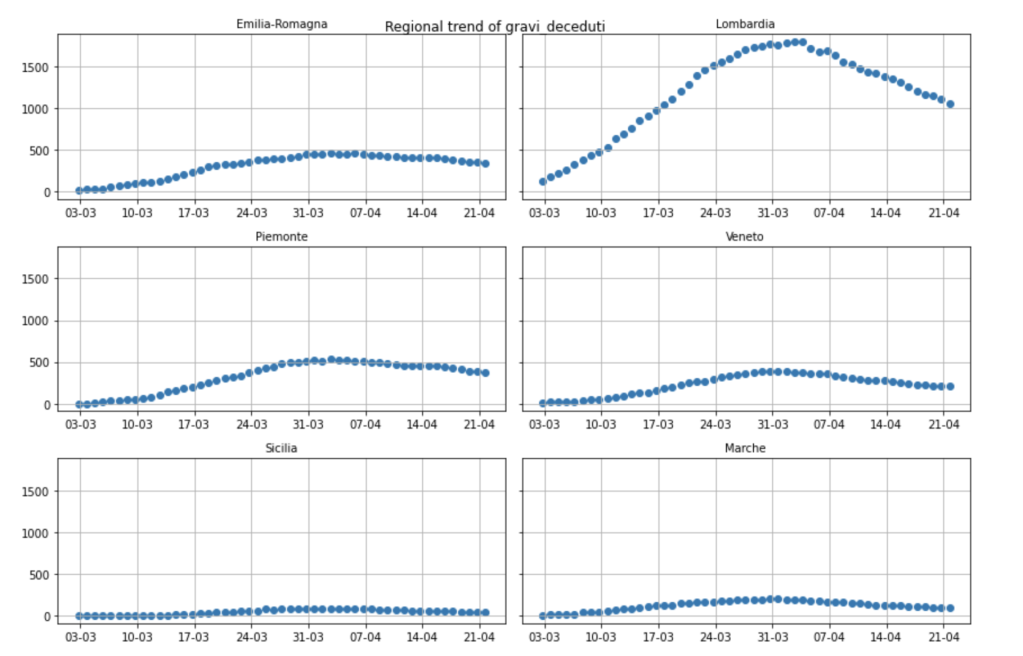
Pl**ot the forecast**



**Regional trends**

The above analysis is treating Italian COVID-19 epidemic on a national basis,. In order to quantify this, we can inspect the regional COVID-19 dataset provided by the Civil Protection, and calculate the proportion of the deaths registered in those regions

If just those 4 regions (out of 20) account for more than 80% of the total deaths, having a single model that predicts the trend for the whole country is a major oversimplification. In fact, the situation varies considerably from region to region, as shown in the image below.



For each region a different model should be applied, in order to better understand the epidemic trends in the different areas of the country. Given that the data on Github.

This Analysis tells us that the COVID-19 epidemic will end in Italy in one and two year, if the current trend is maintained over time.This data set and result shows that Other Country Like India has a low number of cases per one lakh people and this is significant if one considers the availability of resources and population density in India.An important element in the data here is the doubling time. Doubling time shows the number of days it takes for total confirmed cases to double and is a good indicator to see the growth of COVID-19.**The doubling rate is way higher India(cases double every 14 days) ,**United states (35), Russia (20), UK (35), Spain (56), Italy (55), and Germany (54). The only exception to this is Brazil with a doubling time of just 13 days.